On the Identity Problem for the Special Linear Group and the Heisenberg Group

Sang-Ki Ko[1], Reino Niskanen[2], and Igor Potapov[3]

1 **Department of Computer Science, University of Liverpool, United Kingdom**
sangkiko@liverpool.ac.uk
2 **Department of Computer Science, University of Liverpool, United Kingdom**
r.niskanen@liverpool.ac.uk
3 **Department of Computer Science, University of Liverpool, United Kingdom**
potapov@liverpool.ac.uk

Abstract

We study the Identity problem for matrix semigroups. The Identity problem is to decide whether there exists the identity matrix in the given matrix semigroup. It has been recently shown that the Identity problem is NP-complete for a matrix semigroup generated by matrices from the Special Linear Group $\mathrm{SL}(2,\mathbb{Z})$ and undecidable for matrices from $\mathrm{SL}(4,\mathbb{Z})$. In this paper we are filling the gap between these results. First we improve the previous undecidability result that holds for a matrix semigroup generated by 48 4×4 matrices, reducing the bound 48 to 9 and provide a new reduction technique by exploiting the anti-diagonal entries. Next, we analyse the Special Linear Group $\mathrm{SL}(3,\mathbb{Z})$ and show that there is no embedding from a set of pairs over a semigroup alphabet to any matrix semigroup in $\mathrm{SL}(3,\mathbb{Z})$ and therefore there is no embedding from a set of pairs over a group alphabet to any matrix semigroup in $\mathbb{Z}^{3\times 3}$. This implies that any direct encoding of the Post Correspondence Problem or the Identity Correspondence Problem cannot be successfully applied to prove the undecidability of the Identity problem in dimension three over integers. Finally, we consider a well-known subgroup of $\mathrm{SL}(3,\mathbb{Q})$, the Heisenberg group $\mathrm{H}(3,\mathbb{Q})$, which consists of upper-triangular matrices over rationals with determinant 1 in dimension three. We show that the Identity problem for a matrix semigroup generated by matrices from $\mathrm{H}(3,\mathbb{Q})$ is decidable in polynomial time. As the Identity problem is computationally equivalent to the Group problem (i.e., to decide whether a semigroup is a group), all above results hold for the Group problem as well.

1998 ACM Subject Classification F.2.1 Numerical Algorithms and Problems, F.1.1 Models of Computation

Keywords and phrases Matrix semigroups, Identity problem, Special linear group, Heisenberg group

1 Introduction

Many computational problems for matrix semigroups and groups are variants of the reachability problems in infinite state systems. Usually, they are computationally hard starting from dimension two and very often become undecidable from dimension three and four even in the case of integer matrices. The central decision problem in matrix semigroup is the membership problem, which was originally considered by A. Markov in 1947 [14]. Let $S = \langle G \rangle$ be a matrix semigroup finitely generated by a generating set of square matrices G. The *membership problem* is to decide whether or not a given matrix M belongs to the matrix semigroup S. By restricting M to be the identity matrix, we call the problems the *Identity problem*. The other classical decision problems for matrix semigroups are the vector reachability, scalar reachability problems, freeness, matrix/vector/scalar ambiguity,

etc. The large variety of undecidability proofs for these problems based on the design of different encodings to embed symbolic computations of Turing machines, Post Correspondence Problem or Minsky machines into numerical representation in matrices and matrix products [10, 15, 16]. Many nontrivial algorithms for solving decision problems on matrix semigroups are developed, when considering matrices under different constraints like the dimension of matrices, number of matrices in the generator set, or considering specific subclasses of matrices: e.g., the general class of commutative matrices [1], non-commutative case of row-monomial matrices [12] or various subclasses of 2×2 matrix semigroups generated by non-singular integer matrices [18], upper-triangular integer matrices [11], matrices from the special linear group [2, 7], etc.

In some cases, the constraints on the dimension of matrices may not be so effective in the process of finding decidable fragments. For example, let us consider matrices over hypercomplex numbers, such as quaternions, for which associativity still holds, but there is no more commutativity that we have in case of integer, rational or complex numbers. It is known that vast majority of the decision problems for matrices over quaternions are undecidable in dimension two and open in dimension one [3]. The later case can be seen as a specific fragment of 2×2 matrices over complex numbers.

It is believed that the problems for 2×2 matrices over $\mathbb{Z}, \mathbb{Q}, \mathbb{C}$ could be decidable as it was previously shown that there is no injective morphism form a set of pairs over a semigroup alphabet into any matrix semigroup in $\mathbb{C}^{2\times2}$ [6], so any direct embedding of Turing machine computations into $\mathbb{C}^{2\times2}$ does not exist.

In this paper, we have a significant breakthrough result which expands a potential horizon in decidability area for matrix semigoups. We show, that there is no injective morphism from a set of pairs over a semigroup alphabet into any matrix semigroup in $\mathrm{SL}(3,\mathbb{Z})$ or the Heisenberg group $\mathrm{H}(3,\mathbb{C})$, and also there is no embedding from a set of pairs over a group alphabet to any matrix semigroup in $\mathbb{Z}^{3\times3}$. This implies that any direct encoding of the Post Correspondence Problem (defined on a semigroup alphabet) or the Identity Correspondence Problem (defined on a group alphabet) cannot be successfully applied to prove the undecidability of reachability problems for these classes in a direct way.

Also, we study the Identity problem for matrix semigroups, which is known to be NP-complete for matrix semigroups generated by matrices from the Special Linear Group $\mathrm{SL}(2,\mathbb{Z})$ [2] and undecidable for matrices from $\mathrm{SL}(4,\mathbb{Z})$ [4]. As the Identity problem is computationally equivalent to the Group problem (i.e., to decide whether a semigroup is a group), almost all hardness and decidability results which holds for Identity problem can be translated into the Group problem in these classes of matrix semigroups.

The other major result of the paper is the decidability of the Identity problem for a subgroup of $\mathrm{SL}(3,\mathbb{Q})$, which consists of upper-triangular matrices with ones on the main diagonal, known as the Heisenberg group $\mathrm{H}(3,\mathbb{Q})$. We prove that the Identity problem is decidable for matrix semigroups generated by matrices from $\mathrm{H}(3,\mathbb{Q})$ in polynomial time. The main part of the proof is showing that if there exists a sequence of matrices such that the resulting matrix is $\left(\begin{smallmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, for some $c \in \mathbb{Z}$, then there exists a permutation of the sequence such that $-c$ is in the upper corner. Pair of such matrices, if exists, can be used for generating the identity.

Finally, we improve the undecidability result on the Identity problem. In [4], it was shown that the identity problem is undecidable for semigroups consisting of 48 4×4 integer matrices with determinant 1, that is for $\mathrm{SL}(4,\mathbb{Z})$. The idea of the proof is to encode pair of words on the main diagonal. We improve the bound on the number of matrices in the generating set from 48 to 9 and provide a new reduction technique by exploiting the anti-diagonal entries.

2 Preliminaries

Basic definitions. A *semigroup* is a set equipped with an associative binary operation. Let S be a semigroup and X be a subset of S. We say that a semigroup S is *generated* by a subset X of S if each element of S can be expressed as a composition of elements of X. Then, we call X the *generating set* of S. Then, X is a *code* if and only if every element of S has a unique factorization over X. A semigroup S is *free* if there exists a subset $X \subseteq S$ which is a code and $S = X^+$.

Given an alphabet $\Sigma = \{1, 2, \ldots, m\}$, a word w is an element of Σ^*. For a letter $a \in \Sigma$, we denote by $\overline{a}$ the inverse letter of a such that $a\overline{a} = \varepsilon$ where ε is the empty word.

Special linear group, Heisenberg group and its properties. The special linear group is $\mathrm{SL}(n, \mathbb{K}) = \{M \in \mathbb{K}^{n\times n} \mid \det(M) = 1\}$, where $\mathbb{K} = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \ldots$. The Heisenberg group $\mathrm{H}(3, \mathbb{R})$ is formed by the 3×3 real matrices of the form:

$$M = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a, b, c \in \mathbb{R}.$$

It is easy to see that the Heisenberg group is a non-commutative subgroup of $\mathrm{SL}(3, \mathbb{R})$ since the determinant of the matrices in the Heisenberg group is 1. We can consider the Heisenberg group as a set of all triples with the following group law:

$$(a_1, b_1, c_1)(a_2, b_2, c_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2 + a_1b_2). \tag{1}$$

For a matrix M we denote by $\psi(M)$ the triple $(a, b, c) \in \mathbb{R}^3$ which corresponds to the upper-triangular coordinates of M.

Let M be a matrix in $\mathrm{H}(3, \mathbb{R})$ such that $\psi(M) = (a, b, c)$. We define the *superdiagonal vector* of M to be $\vec{v}(M) = (a, b)$. Given two vectors $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$, the *cross product* of $\vec{u}$ and $\vec{v}$ is defined as $\vec{u} \times \vec{v} = u_1v_2 - u_2v_1$. Any two vectors are said to be *parallel* if the cross product is zero.

Group alphabet encodings. It is well-known that $\{a^ib\overline{a}^i \mid i \geq 1\}$ freely generates a free subgroup of the free group $\langle a, b\rangle$ [5] and that the matrices $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ freely generate a free subgroup of $\mathrm{SL}(2, \mathbb{Z})$ [13].

Let $\Sigma = \{z_1, z_2, \ldots, z_\ell\}$ be a group alphabet and $\Sigma_2 = \{a, b, \overline{a}, \overline{b}\}$ be a binary group alphabet. Define the mapping $\alpha : \Sigma \to \Sigma_2^*$ by: $\alpha(z_i) = a^ib\overline{a}^i, \alpha(\overline{z_i}) = a^i\overline{b}\overline{a}^i$, where $1 \leq i \leq \ell$. It is easy to see that α is a monomorphism. Note that α can be extended to domain Σ^* in the usual way. We also define a monomorphism $f : \Sigma_2^* \to \mathbb{Z}^{2\times 2}$ as $f(a) = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$, $f(\overline{a}) = \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$, $f(b) = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ and $f(\overline{b}) = \left(\begin{smallmatrix} 1 & 0 \\ -2 & 1 \end{smallmatrix}\right)$.

The composition of two monomorphisms α and f gives us the embedding from an arbitrary group alphabet into the special linear group $\mathrm{SL}(2, \mathbb{Z})$.

Post correspondence problem (PCP) and its variants. The PCP is a famous undecidable problem introduced by Post in 1946 [17]. Let $\Sigma = \{a, b\}$ be a binary alphabet and $P = \{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\} \subseteq \Sigma^* \times \Sigma^*$ be a set of pairs of words where $n \geq 2$. Then, the PCP is to determine if there exists a finite sequence of indices $\ell_1, \ell_2, \ldots, \ell_k$ with each $1 \leq \ell_i \leq n$ such that: $u_{\ell_1}u_{\ell_2} \cdots u_{\ell_k} = v_{\ell_1}v_{\ell_2} \cdots v_{\ell_k}$.

Currently, the best known undecidability bound for the PCP is $n = 5$ [15]. We denote the minimal number of pairs for which the PCP is undecidable by n_p and use it to describe

the other undecidability bounds presented in the paper. There exists a restricted variant of PCP called the *Restricted PCP*, where the solution starts with (u_1, v_1), ends with (u_n, v_n) and these pairs are used exactly once.The undecidability bound for the Restricted PCP is denoted by n_r and currently, $n_r = 7$.

The *Identity Correspondence Problem (ICP)* [4] is another variant of the PCP which asks whether a finite set of pairs of words over a group alphabet can generate the identity pair by a sequence of concatenations. Let $\Sigma = \{a, b\}$ be a binary alphabet and $\Pi = \{(s_1, t_1), (s_2, t_2), \ldots, (s_m, t_m)\} \subseteq \mathrm{FG}(\Sigma) \times \mathrm{FG}(\Sigma)$. Then, the ICP asks if there exists a nonempty finite sequence of indices $\ell_1, \ell_2, \ldots, \ell_k$ where $1 \leq \ell_i \leq m$ such that $s_{\ell_1} s_{\ell_2} \cdots s_{\ell_k} = t_{\ell_1} t_{\ell_2} \cdots t_{\ell_k} = \varepsilon$, where ε is the empty word. Bell and Potapov [4] proved that the ICP is undecidable by a constructive reduction from the Restricted PCP and showed that the undecidability bound for the ICP is $8(n_r - 1)$ (currently, 48).

3 The Identity Problem in Matrix Semigroups in Dimension Four

▶ **Theorem 1.** *Given a semigroup S generated by a fixed number m of 4×4 integer matrices, determining whether the identity matrix belongs to S is undecidable. This holds for $m = n_p + 4$ (currently, $n_p = 5$).*

Proof. We shall use an encoding to embed an instance of the PCP into a set of 4×4 integer matrices. We use the composition of two monomorphisms α and f introduced in Section 2 to encode a set of pairs of words over an arbitrary group alphabet into a set of 4×4 integer matrices in $\mathrm{SL}(4, \mathbb{Z})$.

Let $P = \{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\} \subseteq \Sigma^* \times \Sigma^*$ be an instance of the PCP. Without loss of generality, we can assume that the first pair of words of the solution is (u_1, v_1). We define the alphabet $\Gamma = \Sigma \cup \Sigma_B$, where $\Sigma = \{a, b\}$ is the alphabet used in the instance of the PCP and $\Sigma_B = \{q_0, q_1, p_0, p_1\}$ is the alphabet for encoding the states of the automaton described in Fig. 1.

Then, we define the following sets of words $W_1 \cup W_2 \subseteq \mathrm{FG}(\Gamma) \times \mathrm{FG}(\Gamma)$, where

- $W_1 = \left\{ \frac{q_0}{p_0} \cdot \frac{a}{a} \cdot \frac{\overline{q_0}}{\overline{p_0}},\ \frac{q_0}{p_0} \cdot \frac{b}{b} \cdot \frac{\overline{q_0}}{\overline{p_0}} \mid a, b \in \Sigma,\ \ q_0, p_0 \in \Sigma_B \right\}$ and
- $W_2 = \left\{ \frac{q_0}{p_0} \cdot \frac{\overline{u_1}}{\overline{v_1}} \cdot \frac{\overline{q_1}}{\overline{p_1}},\ \frac{q_1}{p_1} \cdot \frac{\overline{u_i}}{\overline{v_i}} \cdot \frac{\overline{q_1}}{\overline{p_1}} \mid 1 \leq i \leq n,\ \ (u_i, v_i) \in P,\ \ q_0, q_1, p_0, p_1 \in \Sigma_B \right\}$.

First we prove that $(q_0\overline{q_1}, p_0\overline{p_1}) \in (W_1 \cup W_2)^*$ if and only if the PCP has a solution. It is easy to see that any pair of words in W_1^+ is of the form $(q_0 w \overline{q_0}, p_0 w \overline{p_0})$ for $w \in \Sigma^+$. Then, there exists a pair of words in W_2^* of the form $(q_0 \overline{w}\, \overline{q_1}, p_0 \overline{w}\, \overline{p_1})$ for some word $w \in \Sigma^*$ if and only if the PCP has a solution. Therefore, the pair of words $(q_0\overline{q_1}, p_0\overline{p_1})$ can be constructed by concatenating pairs of words in W_1 and W_2 if and only if the PCP has a solution.

For each pair of words $(u, v) \in W_1 \cup W_2$, we define a matrix $A_{u,v}$ as follows:

$$\begin{pmatrix} f(\alpha(u)) & \mathbf{0}_2 \\ \mathbf{0}_2 & f(\alpha(v)) \end{pmatrix} \in \mathrm{SL}(4, \mathbb{Z}),$$

where $\mathbf{0}_2$ denotes the zero matrix in $\mathbb{Z}^{2\times 2}$. Moreover, we define the following matrix

$$B_{q_1\overline{q_0}, p_1\overline{p_0}} = \begin{pmatrix} \mathbf{0}_2 & f(\alpha(q_1\overline{q_0})) \\ f(\alpha(p_1\overline{p_0})) & \mathbf{0}_2 \end{pmatrix} \in \mathrm{SL}(4, \mathbb{Z}).$$

Let S be a matrix semigroup generated by the set $\{A_{u,v}, B_{q_1\overline{q_0}, p_1\overline{p_0}} \mid (u, v) \in W_1 \cup W_2\}$. We already know that the pair $(q_0\overline{q_1}, p_0\overline{p_1})$ of words can be generated by concatenating

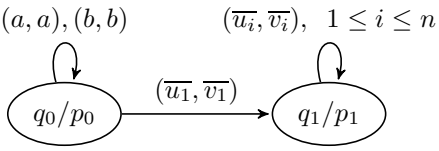


Figure 1 Structure of the automaton encoded by the border letters in Γ_B

words in W_1 and W_2 if and only if the PCP has a solution. The matrix semigroup S has the corresponding matrix $A_{q_0\overline{q_1},p_0\overline{p_1}}$ and thus,

$$\begin{pmatrix} f(\alpha(q_0\overline{q_1})) & \mathbf{0}_2 \\ \mathbf{0}_2 & f(\alpha(p_0\overline{p_1})) \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & f(\alpha(q_1\overline{q_0})) \\ f(\alpha(p_1\overline{p_0})) & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_2 & f(\alpha(\varepsilon)) \\ f(\alpha(\varepsilon)) & \mathbf{0}_2 \end{pmatrix} \in S.$$

Let us denote the identity matrix in $\mathbb{Z}^{n\times n}$ by $\mathbf{I}_n$. Then, we see that the identity matrix $\mathbf{I}_4$ exists in the semigroup S as follows:

$$\begin{pmatrix} \mathbf{0}_2 & f(\alpha(\varepsilon)) \\ f(\alpha(\varepsilon)) & \mathbf{0}_2 \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & f(\alpha(\varepsilon)) \\ f(\alpha(\varepsilon)) & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} f(\alpha(\varepsilon)) & \mathbf{0}_2 \\ \mathbf{0}_2 & f(\alpha(\varepsilon)) \end{pmatrix} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{I}_2 \end{pmatrix} = \mathbf{I}_4 \in S.$$

Now we prove that the identity matrix does not exist in S if the PCP has no solution. It is easy to see that we cannot obtain the identity matrix only by multiplying 'A' matrices since there is no possibility of cancelling every border letter. We need to multiply the matrix $B_{q_1\overline{q_0},p_1\overline{p_0}}$ with a product of 'A' matrices at some point to reach the identity matrix. Note that the matrix $B_{q_1\overline{q_0},p_1\overline{p_0}}$ cannot be the first matrix of the product, followed by the 'A' matrices, because the top right block of $B_{q_1\overline{q_0},p_1\overline{p_0}}$, which corresponds to the first word of the pair, should be multiplied with the bottom right block of 'A' matrix, which corresponds to the second word of the pair. Suppose that the 'A' matrix is of the following form:

$$\begin{pmatrix} f(\alpha(q_0u\overline{q_1})) & \mathbf{0}_2 \\ \mathbf{0}_2 & f(\alpha(p_0v\overline{p_1})) \end{pmatrix}.$$

Since the PCP instance has no solution, either u or v is not the empty word. We multiply $B_{q_1\overline{q_0},p_1\overline{p_0}}$ to the matrix and then obtain the following matrix:

$$\begin{pmatrix} f(\alpha(q_0u\overline{q_1})) & \mathbf{0}_2 \\ \mathbf{0}_2 & f(\alpha(p_0v\overline{p_1})) \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & f(\alpha(q_1\overline{q_0})) \\ f(\alpha(p_1\overline{p_0})) & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0}_2 & f(\alpha(q_0u\overline{q_0})) \\ f(\alpha(p_0v\overline{p_0})) & \mathbf{0}_2 \end{pmatrix}.$$

We can see that either the upper right part or the lower left part cannot be $f(\alpha(\varepsilon))$, which actually corresponds to the identity matrix in $\mathbb{Z}^{2\times 2}$. Now the only possibility of reaching the identity matrix is to multiply matrices which have $\mathrm{SL}(2,\mathbb{Z})$ matrices in the anti-diagonal entries like $B_{q_1\overline{q_0},p_1\overline{p_0}}$. However, we cannot cancel the parts because the top right block (the bottom left block) of the left matrix is multiplied with the bottom left block (the top right block) of the right matrix as follows:

$$\begin{pmatrix} \mathbf{0}_2 & A \\ B & \mathbf{0}_2 \end{pmatrix}\begin{pmatrix} \mathbf{0}_2 & C \\ D & \mathbf{0}_2 \end{pmatrix} = \begin{pmatrix} AD & \mathbf{0}_2 \\ \mathbf{0}_2 & BC \end{pmatrix},$$

where A, B, C and D are matrices in $\mathbb{Z}^{2\times 2}$. As the first word of the pair is encoded in the top right block of the matrix and the second word is encoded in the bottom left block, it is not difficult to see that we cannot cancel the remaining blocks by such multiplications. ◀

Note that the proof above also applies to the special case of the membership problem called the *special diagonal membership problem*, where the task is to determine whether a scalar multiple of the identity matrix exists in a given matrix semigroup. The most recent undecidability bound is shown to be 14 by Halava et al. [10] using the Claus instances of PCP. We further improve the bound to 9 using the same argument as in the proof of Theorem 1. Also note, that in our construction the identity matrix is the only diagonal matrix of the semigroup S, from which we have the following corollary.

▶ **Corollary 2.** *Given a semigroup S generated by a fixed number n of 4×4 integer matrices, determining whether there exists any diagonal matrix in S is undecidable. This holds even for $n = 9$.*

4 Non-existence of Embedding in Dimension Three

Decidability or undecidability of the identity problem for matrix semigroups in dimension three has been open for a long time. Recall that the identity problem has been proven to be NP-complete in dimension two and undecidable in dimension four. It is quite difficult to show the decidability of the problem due to the complicated representation of the special linear group $\mathrm{SL}(3,\mathbb{Z})$ [8, 9] compared to the relatively simple representation of $\mathrm{SL}(2,\mathbb{Z})$. At the same time, it is also possible that the identity problem in dimension three is undecidable but the proof technique should be completely different from the techniques we used in dimension four as we prove that there does not exist an embedding from a set of pairs of words over a semigroup alphabet into the matrix semigroup in $\mathrm{SL}(3,\mathbb{Z})$ and, moreover, from a set of pairs of words over a group alphabet into the matrix semigroup $\mathbb{Z}^{3\times 3}$.

As a first step, we show that there is no embedding from a set of pairs of words over a (semigroup) alphabet Σ to the special linear group $\mathrm{SL}(3,\mathbb{Z})$. It is sufficient to prove the theorem in the case when $\Sigma = \{0,1\}$. In this case, the monoid $S = \Sigma^* \times \Sigma^*$ has a generating set $G = \{(0,\varepsilon),(1,\varepsilon),(\varepsilon,0),(\varepsilon,1),(\varepsilon,\varepsilon)\}$, where ε is the empty word. We simplify the notation by setting $a = (0,\varepsilon), b = (1,\varepsilon), c = (\varepsilon,0), d = (\varepsilon,1)$ and $e = (\varepsilon,\varepsilon)$. It is easy to see that we have the following relations:

$$\begin{aligned} &ac = ca, \quad bc = cb, \quad ad = da, \quad bd = da, \\ &ae = ea, \quad be = eb, \quad ce = ec, \quad de = ed. \end{aligned} \tag{2}$$

In other words, a and b commute with c and d, and a, b, c, d all commute with e. Note that a and b should not commute with each other, and neither should c and d. Let $\phi : S \to \mathrm{SL}(3,\mathbb{Z})$ be an injective morphism and denote $A = \phi(a), B = \phi(b), C = \phi(c), D = \phi(d)$, and $E = \phi(e)$. Our goal is to show that ϕ does not exist. First, we obtain the following restriction for the matrices A, B, C, D:

▶ **Lemma 3.** *If there is an injective morphism $\phi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3,\mathbb{Z})$ and the matrices A, B, C and D correspond to $(0,\varepsilon),(1,\varepsilon),(\varepsilon,0)$ and $(\varepsilon,1)$ respectively, then the Jordan normal form of the matrices A, B, C and D is $\left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Moreover, the matrices have a single eigenvalue.*

Proof. Let ϕ be an injective morphism from S into $\mathrm{SL}(3,\mathbb{C})$.

Let $M, M', N_1, N_2 \in \{A, B, C, D\}$, such that $MM' \neq M'M$, $MN_1 = N_1M$, $MN_2 = N_2M$ and $N_1 \neq N_2$. For example, if $M = A$, then $M' = B$, $N_1 = C$ and $N_2 = D$, or $N_1 = D$ and $N_2 = C$. Since the conjugation by an invertible matrix does not influence the injectivity, we suppose that M is in the Jordan normal form. For a 3×3 matrix, there are six different types of matrices in the Jordan normal form. It is easy to rule out the cases,

where M has more than one eigenvalue because $\det(M) = 1$ and $\mathrm{tr}(M) \in \mathbb{Z}$ implies that all eigenvalues equal to 1. If M has a single eigenvalue, there are three cases. In the first case, $M = \mathbf{I}_3$, which commutes with M'. In the second case, $M = \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. By solving N_1 and N_2 from the relations $MN_1 = N_1M$ and $MN_2 = N_2M$, we see that $N_1N_2 = N_2N_1$. This leaves us with the final case, that is $M = \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. See the appendix for additional details. ◀

Based on the restriction on the Jordan normal form of matrices, we prove that there is no injective morphism from the set of pairs of words over an alphabet Σ into $\mathrm{SL}(3, \mathbb{Z})$.

▶ **Theorem 4.** *There is no injective morphism $\phi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3, \mathbb{Z})$ for any alphabet Σ with at least two elements.*

Proof. Assume to the contrary that there is an injective morphism ϕ from $\Sigma^* \times \Sigma^*$ into $\mathrm{SL}(3, \mathbb{Z})$. Since the conjugation by an invertible matrix does not influence the injectivity, we suppose that the matrix A, which corresponds to the generator $(0, \varepsilon)$, is in the Jordan normal form as proven in Lemma 3. Then, we have the following matrices corresponding to the generators $(0, \varepsilon), (1, \varepsilon), (\varepsilon, 0)$ and $(\varepsilon, 1)$ as follows:

$$A = \phi((0,\varepsilon)) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \phi((1,\varepsilon)) = \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix},$$

$$C = \phi((\varepsilon,0)) = \begin{pmatrix} a_C & b_C & c_C \\ d_C & e_C & f_C \\ g_C & h_C & \ell_C \end{pmatrix}, \text{ and } \qquad D = \phi((\varepsilon,1)) = \begin{pmatrix} a_D & b_D & c_D \\ d_D & e_D & f_D \\ g_D & h_D & \ell_D \end{pmatrix}.$$

Since A and C commute with each other by one of the given relations in (2), it follows that $d_C = g_C = f_C = 0$ and $a_C = e_C$. Therefore, $C = \left(\begin{smallmatrix} a_C & b_C & c_C \\ 0 & a_C & 0 \\ 0 & h_C & \ell_C \end{smallmatrix}\right)$ and $D = \left(\begin{smallmatrix} a_D & b_D & c_D \\ 0 & a_D & 0 \\ 0 & h_D & \ell_D \end{smallmatrix}\right)$. Since C and D are in $\mathrm{SL}(3, \mathbb{Z})$, the determinant of C and D should be 1. Now, the determinant of C is $a_C^2\ell_C$ and hence, $a_C = \ell_C = 1$. Analogously, we can also see that $a_D = \ell_D = 1$. Next we observe that the matrices C and D commute if and only if $c_Ch_D = c_Dh_C$.

We have the three cases where C and D cannot commute (equivalently, $c_Ch_D \neq c_Dh_C$): 1) $c_C = 0$ and $h_C \neq 0$; 2) $c_C \neq 0$ and $h_C \neq 0$; and 3) $c_C \neq 0$ and $h_C = 0$.

In the first case, by solving B from the relations $BC = CB$ and $BD = DB$, we see that $c_B = d_B = f_B = 0$. By Lemma 3, B has a single eigenvalue and therefore $B = \left(\begin{smallmatrix} 1 & b_B & 0 \\ 0 & 1 & 0 \\ g_B & h_B & 1 \end{smallmatrix}\right)$. Now, from the relation $BD = DB$, we have that $b_D = c_D$ and D commutes with C, which is a contradiction.

The remaining cases are analogous and also lead to a contradiction. See the appendix for the details. Since we have examined all possible cases and found contradictions for every case, we can conclude that there is no injective morphism from $\Sigma^* \times \Sigma^*$ into the special linear group $\mathrm{SL}(3, \mathbb{Z})$. ◀

▶ **Corollary 5.** *There is no injective morphism $\phi : \mathrm{FG}(\Sigma) \times \mathrm{FG}(\Sigma) \to \mathbb{Z}^{3\times3}$ for any alphabet Σ with at least two elements.*

Proof. We proceed by contradiction. Assume that there exists such an injective morphism ϕ from the set of pairs of words over a group alphabet to the set of matrices in $\mathbb{Z}^{3\times3}$. Suppose that $A = \phi(a, \varepsilon)$ where $a \in \Sigma$. Then, the inverse matrix A^{-1} corresponding to $(\overline{a}, \varepsilon)$ must be in $\mathbb{Z}^{3\times3}$. This implies that the determinant of A is 1 because otherwise the determinant of A^{-1} becomes a non-integer. ◀

Next, we show that there does not exist an embedding from pair of words into the Heisenberg group over complex numbers. If we consider the discrete Heisenberg group $\mathrm{H}(3,\mathbb{Z})$, then the result follows from Theorem 4.

▶ **Theorem 6.** *There is no injective morphism $\varphi : \Sigma^* \times \Sigma^* \to \mathrm{H}(3,\mathbb{C})$ for any alphabet Σ with at least two elements.*

Proof. Assume to the contrary that there is an injective morphism φ from $\Sigma^* \times \Sigma^*$ into $\mathrm{H}(3,\mathbb{C})$. Using the notations and relations of (2), we set $\varphi(a) = A$, $\varphi(b) = B$, $\varphi(c) = C$, $\varphi(d) = D$, $\varphi(e) = E$ for some matrices $A, B, C, D, E \in \mathrm{H}(3,\mathbb{C})$. It is easy to see that two matrices $M, N \in \mathrm{H}(3,\mathbb{C})$ commute if and only if $\vec{v}(M) \times \vec{v}(N) = 0$. Denote $\vec{v}(A) = (a_1, a_2)$ and $\vec{v}(B), \vec{v}(C), \vec{v}(D), \vec{v}(E)$ are denoted analogously. From the relations (2), it follows that

$$a_1c_2 = c_1a_2, \quad a_1d_2 = d_1a_2, \quad b_1c_2 = c_1b_2, \quad b_1d_2 = d_1b_2, \quad a_1b_2 \neq b_1a_2, \quad c_1d_2 \neq d_1c_2.$$

A simple calculation shows that the relations cannot hold, which contradicts our assumption that φ exists. ◀

5 Decidability of the Identity Problem in the Heisenberg group

As the decidability status of the identity problem in dimension three is unknown in general, we look for a subclass of $\mathrm{SL}(3,\mathbb{Z})$ for which the identity problem could be decidable. The Heisenberg group is an interesting subgroup of $\mathrm{SL}(3,\mathbb{Z})$ which is useful in the description of one-dimensional quantum mechanical systems.

Considering the multiplication of matrices in the Heisenberg group, we simply add up the numbers in the two superdiagonal coordinates in a commutative way. In this section, we prove that the identity problem for matrix semigroups in the Heisenberg group over rationals by analyzing the behaviour of multiplications especially in the upper-right coordinate of matrices. First, we establish the following property of the Heisenberg group.

▶ **Lemma 7.** *Let $G = \{M_1, M_2, \ldots, M_n\} \subseteq \mathrm{H}(3,\mathbb{R})$ be a set of matrices from the Heisenberg group such that superdiagonal vectors of matrices are pairwise parallel. If there exists a sequence of matrices $M = M_{i_1}M_{i_2}\cdots M_{i_k}$, where $i_j \in [1,n]$ for all $1 \leq j \leq k$, such that $\psi(M) = (0,0,c)$ for some $c \in \mathbb{R}$, then, $c = \sum_{j=1}^{k}\left(c_{i_j} - \frac{x}{2}a_{i_j}^2\right)$.*

Proof. Let M and M' be two matrices from the Heisenberg group. The superdiagonal vectors of M and M' are parallel if and only if $\vec{v}(M) = (a,b)$ and $\vec{v}(M') = (ya, yb)$ for some $y \in \mathbb{R}$. Then, the multiplication of such matrices is always commutative as follows:

$$MM' = \begin{pmatrix} 1 & a+ay & c+c'+aby \\ 0 & 1 & b+by \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & ay+a & c'+c+ayb \\ 0 & 1 & by+b \\ 0 & 0 & 1 \end{pmatrix} = M'M.$$

Now consider the sequence $M_{i_1}M_{i_2}\cdots M_{i_k}$ and let $\psi(M_i) = (a_i, b_i, c_i)$ for each $i \in [1,n]$. Since $\frac{b_i}{a_i}$ is equivalent for all $i \in [1,n]$, let us denote $b_i = a_ix$. From the multiplication of matrices, we have the following equation:

$$\begin{aligned} c &= \sum_{j=1}^{k} c_{i_j} + \sum_{\ell=1}^{k-1}\left(\sum_{j=1}^{\ell} a_{i_j}\right)a_{i_{\ell+1}}x = \sum_{j=1}^{k} c_{i_j} + \frac{1}{2}\left(\sum_{\ell=1}^{k}\sum_{j=1}^{k} a_{i_\ell}a_{i_j}x - \sum_{j=1}^{k} a_{i_j}^2 x\right) \\ &= \sum_{j=1}^{k} c_{i_j} - \frac{1}{2}\sum_{j=1}^{k} a_{i_j}^2 x = \sum_{j=1}^{k}(c_{i_j} - \frac{x}{2}a_{i_j}^2). \end{aligned}$$

From the above equation, we prove the statement claimed in the lemma. Moreover, due to the commutativity of multiplication, the value c does not change even if we change the order of multiplicands. ◀

It is worth mentioning that the identity problem in the Heisenberg group is decidable if any two matrices have pairwise parallel superdiagonal vectors since now the problem reduces to solving a system of two Diophantine equations. Hence, it remains to consider the case when there exist two matrices with non-parallel superdiagonal vectors in the sequence generating the identity matrix. In the following, we prove that the identity matrix is always constructible if we can construct any matrix with the zero superdiagonal vector by using matrices with non-parallel superdiagonal vectors.

▶ **Lemma 8.** *Let $S = \langle M_1, M_2, \ldots, M_n \rangle \subseteq \mathrm{H}(3, \mathbb{Q})$ be a finitely generated matrix semigroup. Then, the identity matrix exists in S if there exists a sequence of matrices $M_{i_1} M_{i_2} \cdots M_{i_k}$, where $i_j \in [1, n]$ for all $1 \leq j \leq k$, satisfying the following properties:*

1. *$\psi(M_{i_1} M_{i_2} \cdots M_{i_k}) = (0, 0, c)$ for some $c \in \mathbb{Q}$, and*
2. *$\vec{v}(M_{i_{j_1}})$ and $\vec{v}(M_{i_{j_2}})$ are not parallel for some $j_1, j_2 \in [1, k]$.*

Proof. Let $M = M_{i_1} M_{i_2} \cdots M_{i_k}$ and $\psi(M) = (0, 0, c)$ for some $c \in \mathbb{Q}$. It is obvious that the identity matrix is in S if $c = 0$. Hence we assume that $c > 0$ as the case of $c < 0$ is symmetric.

Given that M_i is the ith generator and $\psi(M_i) = (a_i, b_i, c_i)$, we have $\sum_{j=1}^{k} a_{i_j} = 0$ and $\sum_{j=1}^{k} b_{i_j} = 0$. Since $c > 0$, the following also holds:

$$c = \sum_{i=1}^{k-1} \sum_{j=1}^{i} a_j b_{i+1} + \sum_{j=1}^{k} c_{i_j} > 0. \tag{3}$$

If the matrix semigroup S in $\mathrm{H}(3, \mathbb{Q})$ has two different matrices M_1 and M_2 such that $\psi(M_1) = (0, 0, c_1)$ and $\psi(M_2) = (0, 0, c_2)$ and $c_1 c_2 < 0$, then the identity matrix should exist in S. Let $\psi(M_1) = (0, 0, \frac{p_1}{q_1})$ and $\psi(M_2) = (0, 0, \frac{p_2}{q_2})$ where $p_1, q_1, q_2 \in \mathbb{Z}$ are positive and $p_2 \in \mathbb{Z}$ is negative. Then, it is easy to see that the matrix $M_1^{-q_1 p_2} M_2^{q_2 p_1}$ exists in S such that $\psi(M_1^{-q_1 p_2} M_2^{q_2 p_1}) = (0, 0, 0)$.

Now we will prove that if S contains a matrix M such that $\psi(M) = (0, 0, c)$ where $c > 0$, then there also exists a matrix M' such that $\psi(M') = (0, 0, c')$ where $c' < 0$.

First, we classify the matrices into four types as follows. A matrix with a superdiagonal vector (a, b) is classified as

1. the $(+, +)$-type if $a, b > 0$,
2. the $(+, -)$-type if $a \geq 0$ and $b \leq 0$,
3. the $(-, -)$-type if $a, b \leq 0$, and
4. the $(-, +)$-type if $a < 0$ and $b > 0$.

Let $G = \{M_1, M_2, \ldots, M_n\}$ be the generating set of the matrix semigroup S. Then, $G = G_{(+,+)} \sqcup G_{(+,-)} \sqcup G_{(-,-)} \sqcup G_{(-,+)}$ such that $G_{(\xi_1, \xi_2)}$ is the set of matrices of the (ξ_1, ξ_2)-type where $\xi_1, \xi_2 \in \{+, -\}$.

The main idea of the proof is to generate a matrix M' such that $\psi(M') = (0, 0, c')$ for some $c' < 0$ by using the sequence $M = M_{i_1} M_{i_2} \cdots M_{i_k}$ multiple times. Note that any permutation of the sequence generating the matrix M such that $\psi(M) = (0, 0, c)$ still generates matrices M' such that $\psi(M') = (0, 0, c')$ since the multiplication of matrices changes the front two coordinates in a commutative way. Moreover, we can still obtain

matrices M'' such that $\psi(M'') = (0, 0, c'')$ for some $c'' \in \mathbb{Q}$ if we shuffle two different permutations of the sequence by the same reason.

We illustrate the idea with the following example. Let $\{M_i \mid 1 \leq i \leq 4\} \subseteq G_{(+,+)}$, $\{M_i \mid 5 \leq i \leq 7\} \subseteq G_{(+,-)}$, $\{M_i \mid 8 \leq i \leq 9\} \subseteq G_{(-,-)}$, and $\{M_i \mid 10 \leq i \leq 13\} \subseteq G_{(-,+)}$. Then, assume that $M_1 M_2 \cdots M_{13} = \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, where x is computed by (3). As we mentioned above, x changes if we change the order of multiplications. See Fig. 2 for example. In this example, we first multiply $(+,+)$-type matrices and accumulate the values in the superdiagonal entries since these matrices have positive values in the entries. Indeed, the blue dotted area implies the value we add to the upper-right corner by multiplying such matrices. Then, we multiply $(+,-)$-type matrices and still increase the 'a'-value. The 'b'-values in $(+,-)$-type matrices are negative thus, the red lined area is subtracted from the upper-right corner. We still subtract by multiplying $(-,-)$-type matrices since the accumulated 'a'-value is still positive and 'b'-values are negative. Then, we finish the multiplication by adding exactly the last blue dotted area to the upper-right corner. It is easy to see that the total subtracted value is larger than the total added value.

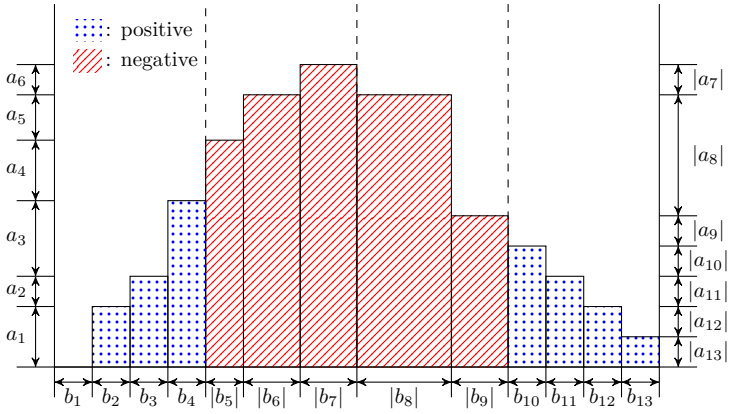


Figure 2 The histogram describes how the upper-right corner of $M_1 M_2 \cdots M_{13}$ is computed by multiplications.

However, we cannot guarantee that x is negative since $\sum_{i=1}^{13} c_i$ could be larger than the contribution from the superdiagonal entries. This is why we need to copy the sequence of matrices generating the matrix corresponding to the triple $(0, 0, c)$ for some $c \in \mathbb{Z}$. In Fig. 3, we describe an example where we duplicate the sequence eight times and shuffle and permute them in order to minimize the value in the upper-right corner. Now the lengths of both axes are m ($m = 8$ in Fig. 3) times larger than before and it follows that the area also grows quadratically in m. Since the summation $m \cdot \sum_{i=1}^{13} c_i$ grows linearly in m, we have $x < 0$ when m is large enough.

From the sequence $M_{i_1} M_{i_2} \cdots M_{i_k}$, we obtain four multisets $S_{(\xi_1,\xi_2)}$, where $\xi_1, \xi_2 \in \{+,-\}$, such that each multiset $S_{(\xi_1,\xi_2)}$ contains the matrices that appear in the sequence and belong to the set $G_{(\xi_1,\xi_2)}$. For instance, the multiset $S_{(+,+)}$ has two identical matrices corresponding to $(5, 4, 2)$ which obviously belongs to $G_{(+,+)}$ if the matrix appears twice in the sequence $M_{i_1} M_{i_2} \cdots M_{i_k}$ since we allow any multiset to have multiple identical elements.

For each $\xi_1, \xi_2 \in \{+,-\}$, let us define $a_{(\xi_1,\xi_2)}, b_{(\xi_1,\xi_2)}, c_{(\xi_1,\xi_2)}$ such that

$$(a_{(\xi_1,\xi_2)}, b_{(\xi_1,\xi_2)}, c_{(\xi_1,\xi_2)}) = \sum_{M \in S_{(\xi_1,\xi_2)}} \psi(M).$$

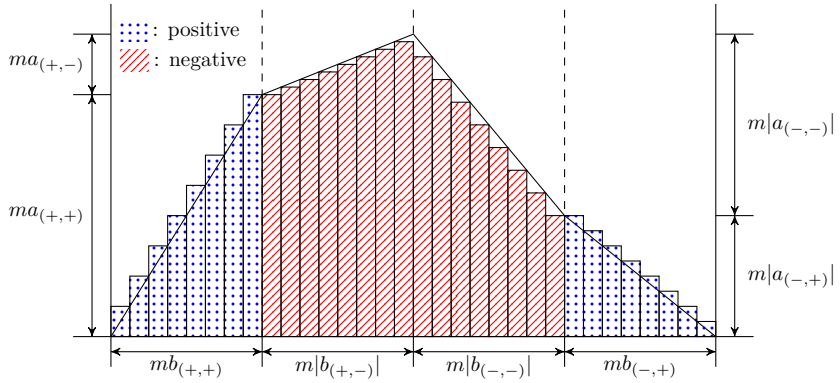


Figure 3 The blue dotted (red lined) area implies the over (under) approximated value which will be added to (subtracted from) the upper-right corner of the final matrix after multiplications of matrices in the sequence. Here $m = 8$.

In other words, $a_{(\xi_1,\xi_2)}$ ($b_{(\xi_1,\xi_2)}$ and $c_{(\xi_1,\xi_2)}$, respectively) is the sum of the values in the 'a' ('b' and 'c', respectively) coordinate from the matrices in the multiset $S_{(\xi_1,\xi_2)}$.

Now consider a permutation of the sequence $M_{i_1}M_{i_2}\cdots M_{i_k}$, where the first part of the sequence only consists of the $(+,+)$-type matrices, the second part only consists of the $(+,-)$-type matrices, the third part only consists of the $(-,-)$-type matrices, and finally the last part only consists of the $(-,+)$-type matrices.

Let us denote by $M_{(+,+)}$ the matrix which results from the multiplication of the first part. Then, the matrix $M_{(+,+)}$ will be bounded from above by the following matrix on the right-hand side:

$$\prod_{M\in S_{(+,+)}} M = \begin{pmatrix} 1 & a_{(+,+)} & x \\ 0 & 1 & b_{(+,+)} \\ 0 & 0 & 1 \end{pmatrix} < \begin{pmatrix} 1 & a_{(+,+)} & c_{(+,+)} + a_{(+,+)}b_{(+,+)} \\ 0 & 1 & b_{(+,+)} \\ 0 & 0 & 1 \end{pmatrix}, \tag{4}$$

where x is computed by (3). Let us define $M_{(+,-)}$, $M_{(-,-)}$ and $M_{(-,+)}$ analogously.

Now we claim that there exists an integer $m > 0$ such that $M^m_{(+,+)}M^m_{(+,-)}M^m_{(-,-)}M^m_{(-,+)}$ corresponds to the triple $(0, 0, c')$ for some $c' < 0$.

Let us first consider the first part $M^m_{(+,+)}$. It follows from (4) that $\psi(M^m_{(+,+)}) = (ma_{(+,+)}, mb_{(+,+)}, mc_{(+,+)} + z_1)$, where z_1 is bounded from above by the following value:

$$z_1 < \sum_{i=1}^{m} i|a_{(+,+)}||b_{(+,+)}| = \frac{m(m+1)}{2} \cdot |a_{(+,+)}||b_{(+,+)}| = z'_1.$$

Now we multiply $M^m_{(+,+)}$ by the second part $M^m_{(+,-)}$. Then, the resulting matrix corresponds to $(m(a_{(+,+)} + a_{(+,-)}), m(b_{(+,+)} + b_{(+,-)}), m(c_{(+,+)} + c_{(+,-)}) + z_1 - z_2)$, where z_2 is bounded from below by the following value:

$$z_2 > m^2|a_{(+,+)}||b_{(+,-)}| + \sum_{i=1}^{m-1} i|a_{(+,-)}||b_{(+,-)}| = z'_2.$$

Similarly, we compute bounds of z_3 and z_4 that are added to the third component of the triple as a result of multiplying $M^m_{(-,-)}$ and $M^m_{(-,+)}$ as follows:

$$z_3 > m^2|a_{(-,+)}||b_{(-,-)}| + \sum_{i=1}^{m-1} i|a_{(-,-)}||b_{(-,-)}| = z'_3 \quad \text{and}$$

$$z_4 < \sum_{i=1}^{m} i|a_{(-,+)}||b_{(-,+)}| = \frac{m(m+1)}{2} \cdot |a_{(-,+)}||b_{(-,+)}| = z'_4.$$

It is easy to see that $\psi(M^m_{(+,+)}M^m_{(+,-)}M^m_{(-,-)}M^m_{(-,+)}) = (0, 0, mc + z_1 - z_2 - z_3 + z_4)$. From the above inequalities we obtain the following inequality:

$$z_1 - z_2 - z_3 + z_4 < z'_1 - z'_2 - z'_3 + z'_4 = z.$$

It is easy to see that z can be represented as a quadratic equation of m using the above results and then the coefficient of m^2 is always negative if $S_{(\xi_1,\xi_2)} \neq \emptyset$ for all $\xi_1, \xi_2 \in \{+, -\}$.

Since the coefficient of the highest power of the variable is negative, z becomes negative when m is large enough. Therefore, we have a matrix corresponding to the triple $(0, 0, c')$ for some $c' < 0$ as a product of multiplying matrices in the generating set and the identity matrix is also reachable.

There are some subcases where one or more subset from $S_{(+,+)}$, $S_{(-,+)}$, $S_{(+,-)}$, and $S_{(-,-)}$ is empty. As in the previous case, it can be shown that the coefficient of m^2 in z is negative and thus, starting from some m, z is negative. See the appendix for detailed analysis of the cases.

Since we have proven that it is always possible to construct a matrix M' such that $\psi(M') = (0, 0, c')$ for some $c' < 0$ in any case, we complete the proof. ◀

▶ **Theorem 9.** *The identity problem for finitely generated matrix semigroups in the Heisenberg group* $\mathrm{H}(3, \mathbb{Q})$ *is decidable in polynomial time.*

Proof. Let S be the matrix semigroup in $\mathrm{H}(3, \mathbb{Q})$ generated by the set $G = \{M_1, M_2, \ldots, M_n\}$. There are two possible cases of having the identity matrix in the matrix semigroup in $\mathrm{H}(3, \mathbb{Q})$. The first case is that the matrices generating the identity matrix have pairwise parallel superdiagonal vectors.

The first case is decidable by reducing to the problem of solving a system of linear Diophantine equations by Lemma 7. We partition the set G into several disjoint subsets $G_1, G_2, \ldots, G_m$ where each subset contains matrices with parallel superdiagonal vectors. Let $G_i = \{M_{k_1}, \ldots, M_{k_{m_i}}\}$ be one of the subsets containing m_i matrices and $M_{k_j} = \begin{pmatrix} 1 & a_{k_j} & c_{k_j} \\ 0 & 1 & b_{k_j} \\ 0 & 0 & 1 \end{pmatrix}$ for all $k \in [1, m]$ and $j \in [1, m_i]$. By Lemma 7, we can transform the matrix M_{k_j} into the following form: $M_{k_j} = \begin{pmatrix} 1 & a_{k_j} & c_{k_j} - \frac{x}{2}a^2_{k_j} \\ 0 & 1 & a_{k_j}x \\ 0 & 0 & 1 \end{pmatrix}, \quad x \in \mathbb{Q}$.

For each G_i, we solve the system of Diophantine equations $A_i\vec{y_i} = \vec{0}$, where

$$A_i = \begin{pmatrix} a_{k_1} & a_{k_2} & \cdots & a_{k_{m_i}} \\ 2c_{k_1} - xa^2_{k_1} & 2c_{k_2} - xa^2_{k_2} & \cdots & 2c_{k_{m_i}} - xa^2_{k_{m_i}} \end{pmatrix}.$$

It is obvious that the identity matrix is in the matrix semigroup if we have a solution in the system of linear Diophantine equations for any subset G_i.

In the second case, it only remains to enforce the condition that the sequence of matrices generating a matrix with zero superdiagonal vector contains two matrices with non-parallel superdiagonal vectors. Therefore, the problem reduces again to solving at most $O(n^2)$ systems of linear Diophantine equations as we have $O(n^2)$ pairs of matrices with non-parallel superdiagonal vectors in the worst-case. Finally, we conclude the proof by mentioning that the identity problem for matrix semigroups in the Heisenberg group over rationals can be even decided in polynomial time as a system of linear Diophantine equations can be solved in polynomial time when the solution is restricted to natural numbers. ◀

References

1 László Babai, Robert Beals, Jin-yi Cai, Gábor Ivanyos, and Eugene M. Luks. Multiplicative equations over commuting matrices. In *Proceedings of Symposium on Discrete Algorithms 1996*, pages 498–507, 1996.

2 Paul C. Bell, Mika Hirvensalo, and Igor Potapov. The identity problem for matrix semigroups in $\mathrm{SL}(2,\mathbb{Z})$ is NP-complete. In *Proceedings of Symposium on Discrete Algorithms 2017*, 2017. doi:10.1137/1.9781611974782.13.

3 Paul C. Bell and Igor Potapov. Reachability problems in quaternion matrix and rotation semigroups. *Information and Computation*, 206(11):1353–1361, 2008.

4 Paul C. Bell and Igor Potapov. On the undecidability of the identity correspondence problem and its applications for word and matrix semigroups. *International Journal of Foundations of Computer Science*, 21(6):963–978, 2010.

5 Jean-Camille Birget and Stuart W. Margolis. Two-letter group codes that preserve aperiodicity of inverse finite automata. *Semigroup Forum*, 76(1):159–168, 2008.

6 Julien Cassaigne, Tero Harju, and Juhani Karhumäki. On the undecidability of freeness of matrix semigroups. *International Journal of Algebra and Computation*, 09(03n04):295–305, 1999.

7 Christian Choffrut and Juhani Karhumäki. Some decision problems on integer matrices. *RAIRO - Theoretical Informatics and Applications*, 39(1):125–131, 3 2010.

8 Marston Conder, Edmund Robertson, and Peter Williams. Presentations for 3-dimensional special linear groups over integer rings. *Proceedings of the American Mathematical Society*, 115(1):19, may 1992. doi:10.2307/2159559.

9 Marston D. E. Conder. Some unexpected consequences of symmetry computations. In *Symmetries in Graphs, Maps, and Polytopes*, pages 71–79. Springer International Publishing, 2016. doi:10.1007/978-3-319-30451-9_3.

10 Vesa Halava, Tero Harju, and Mika Hirvensalo. Undecidability bounds for integer matrices using Claus instances. *International Journal of Foundations of Computer Science*, 18(05):931–948, 2007.

11 Juha Honkala. A Kraft–McMillan inequality for free semigroups of upper-triangular matrices. *Information and Computation*, 239:216–221, 2014. doi:10.1016/j.ic.2014.09.002.

12 Alexei Lisitsa and Igor Potapov. Membership and reachability problems for row-monomial transformations. In *Proceedings of MFCS 2004*, volume 3153 of *LNCS*, pages 623–634, 2004. doi:10.1007/978-3-540-28629-5_48.

13 Roger C. Lyndon and Paul E. Schupp. *Combinatorial group theory*. Springer, 1977.

14 A Markov. On certain insoluble problems concerning matrices. In *Doklady Akad. Nauk SSSR*, volume 57, pages 539–542, 1947.

15 Turlough Neary. Undecidability in binary tag systems and the post correspondence problem for five pairs of words. In *STACS 2015*, volume 30 of *LIPIcs*, pages 649–661, 2015.

16 Michael S Paterson. Unsolvability in 3 by 3 matrices. *Studies in Applied Mathematics*, 49(1):105, 1970.

17 Emil L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52(4):264–268, 1946.

18 Igor Potapov and Pavel Semukhin. Decidability of the membership problem for 2×2 integer matrices. In *Proceedings of Symposium on Discrete Algorithms 2017*, 2017. doi:10.1137/1.9781611974782.12.

A Appendix

▶ **Lemma 3.** *If there is an injective morphism $\phi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3,\mathbb{Z})$ and the matrices A, B, C and D correspond to $(0,\varepsilon), (1,\varepsilon), (\varepsilon,0)$ and $(\varepsilon,1)$ respectively, then the Jordan normal form of the matrices A, B, C and D is $\left(\begin{smallmatrix}1&1&0\\0&1&0\\0&0&1\end{smallmatrix}\right)$. Moreover, the matrices have a single eigenvalue.*

Proof. Let ϕ be an injective morphism from S into $\mathrm{SL}(3,\mathbb{C})$.

Let $M, M', N_1, N_2 \in \{A, B, C, D\}$, such that $MM' \neq M'M$, $MN_1 = N_1M$, $MN_2 = N_2M$ and $N_1 \neq N_2$. For example, if $M = A$, then $M' = B$, $N_1 = C$ and $N_2 = D$ or $N_1 = D$ and $N_2 = C$.

Since the conjugation by an invertible matrix does not influence the injectivity, we suppose that M is in the Jordan normal form. For a 3×3 matrix, there are six different types of matrices in the Jordan normal form. If M has three different eigenvalues, then

$$M = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}. \tag{5}$$

If M has two eigenvalues, then

$$M = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ or } M = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{pmatrix}. \tag{6}$$

Finally, if M has only one eigenvalue, then

$$M = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \text{ or } M = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \text{ or } M = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}. \tag{7}$$

The first case (5) can be easily ruled out since M only commutes with diagonal matrices. Then, N_1 and N_2 should be commuting with M by the suggested relations and as a result, N_1 and N_2 commute with each other.

Now let us consider the second case (6), where M has two eigenvalues λ and μ. Note that the determinant of M is 1 since $M \in \mathrm{SL}(3,\mathbb{Z})$. Namely, $\det(M) = \lambda\mu^2 = 1$. Moreover the trace of M should be an integer since if L is an invertible square matrix of same order as M, then $\mathrm{tr}(M) = \mathrm{tr}(L^{-1}ML)$. This implies that $2 \times \mu + \lambda$ should be an integer and the only possibility for λ and μ to satisfy the two conditions $\lambda\mu^2 = 1$ and $2\mu + \lambda$ being an integer is $\lambda = \mu = 1$, which is reduced to the following case (7).

Finally, we consider the case (7) where M has only one eigenvalue. If the matrix M is diagonal, it is easy to see that it is not the case since otherwise M commutes with all matrices including M'.

If the matrix M is in the following form $\left(\begin{smallmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{smallmatrix}\right)$ and let $N_1 = \left(\begin{smallmatrix}a&b&c\\d&e&f\\g&h&\ell\end{smallmatrix}\right)$. Now

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & \ell \end{pmatrix} = \begin{pmatrix} d + a\lambda & e + b\lambda & f + c\lambda \\ g + d\lambda & h + e\lambda & \ell + f\lambda \\ g\lambda & h\lambda & \ell\lambda \end{pmatrix}$$

$$= \begin{pmatrix} a\lambda & a + b\lambda & b + c\lambda \\ d\lambda & d + e\lambda & e + f\lambda \\ g\lambda & g + h\lambda & h + \ell\lambda \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & \ell \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Since these matrices are equal, we have that $d = g = h = 0$, $a = e = \ell$ and $b = f$. Similar calculation gives us $N_2 = \left(\begin{smallmatrix} a' & b' & c' \\ 0 & a' & b' \\ 0 & 0 & a' \end{smallmatrix}\right)$ and now matrices N_1 and N_2 commute as follows:

$$\begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix} \begin{pmatrix} a' & b' & c' \\ 0 & a' & b' \\ 0 & 0 & a' \end{pmatrix} = \begin{pmatrix} aa' & ab'+ba' & ac'+bb'ca' \\ 0 & aa' & ab'+ba' \\ 0 & 0 & aa' \end{pmatrix} = \begin{pmatrix} a' & b' & c' \\ 0 & a' & b' \\ 0 & 0 & a' \end{pmatrix} \begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix}.$$

Now the only possibility for M is the following form:

$$M = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

where λ is the single eigenvalue of M. Since $\det(M) = \lambda^3 = 1$, λ can be one of cube roots of unity: $1, -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $-\frac{1}{2} - \frac{\sqrt{3}}{2}i$. Among these numbers, the latter two cannot be chosen to be λ since $\mathrm{tr}(M) = 3\lambda$ cannot be an integer. ◀

▶ **Theorem 4.** *There is no injective morphism $\phi : \Sigma^* \times \Sigma^* \to \mathrm{SL}(3, \mathbb{Z})$ for any alphabet Σ with at least two elements.*

Proof. Assume to the contrary that there is an injective morphism ϕ from $\Sigma^* \times \Sigma^*$ into $\mathrm{SL}(3, \mathbb{Z})$. Since the conjugation by an invertible matrix does not influence the injectivity, we suppose that the matrix A, which corresponds to the generator $(0, \varepsilon)$, is in the Jordan normal form as proven in Lemma 3. Then, we have the following matrices corresponding to the generators $(0, \varepsilon), (1, \varepsilon), (\varepsilon, 0)$ and $(\varepsilon, 1)$ as follows:

$$A = \phi((0,\varepsilon)) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \phi((1,\varepsilon)) = \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix},$$

$$C = \phi((\varepsilon,0)) = \begin{pmatrix} a_C & b_C & c_C \\ d_C & e_C & f_C \\ g_C & h_C & \ell_C \end{pmatrix}, \text{ and } \qquad D = \phi((\varepsilon,1)) = \begin{pmatrix} a_D & b_D & c_D \\ d_D & e_D & f_D \\ g_D & h_D & \ell_D \end{pmatrix}.$$

Since A and C commute with each other by one of the given relations in (2), we have

$$AC = \begin{pmatrix} a_C + d_C & b_C + e_C & c_C + f_C \\ d_C & e_C & f_C \\ g_C & h_C & \ell_C \end{pmatrix} = \begin{pmatrix} a_C & a_C + b_C & c_C \\ d_C & d_C + e_C & f_C \\ g_C & g_C + h_C & \ell_C \end{pmatrix} = CA.$$

It is easy to see that $d_C = g_C = f_C = 0$ and $a_C = e_C$. Therefore, C and D should be

$$C = \begin{pmatrix} a_C & b_C & c_C \\ 0 & a_C & 0 \\ 0 & h_C & \ell_C \end{pmatrix} \text{ and } D = \begin{pmatrix} a_D & b_D & c_D \\ 0 & a_D & 0 \\ 0 & h_D & \ell_D \end{pmatrix}.$$

Since C and D are in $\mathrm{SL}(3, \mathbb{Z})$, the determinant of C and D should be 1. Now, the determinant of C is $a_C^2 \ell_C$ and hence, $a_C = \ell_C = 1$. Analogously, we can also see that $a_D = \ell_D = 1$. Next we observe that the matrices commute if and only if $c_C h_D = c_D h_C$:

$$CD = \begin{pmatrix} 1 & b_C & c_C \\ 0 & 1 & 0 \\ 0 & h_C & 1 \end{pmatrix} \begin{pmatrix} 1 & b_D & c_D \\ 0 & 1 & 0 \\ 0 & h_D & 1 \end{pmatrix} = \begin{pmatrix} 1 & b_C + b_D + c_C h_D & c_C + c_D \\ 0 & 1 & 0 \\ 0 & h_C + h_D & 1 \end{pmatrix}.$$

We have the three cases where C and D cannot commute (equivalently, $c_C h_D \neq c_D h_C$): 1) $c_C = 0$ and $h_C \neq 0$; 2) $c_C \neq 0$ and $h_C \neq 0$; and 3) $c_C \neq 0$ and $h_C = 0$.

Let us examine the three cases as follows:

Case 1 ($c_C = 0$ and $h_C \neq 0$). We know that c_D is also non-zero because otherwise C and D commute with each other since $c_C h_D = c_D h_C = 0$. We have the following calculations:

$$BC = \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix} \begin{pmatrix} 1 & b_C & 0 \\ 0 & 1 & 0 \\ 0 & h_C & 1 \end{pmatrix} = \begin{pmatrix} a_B & a_B b_C + b_B + c_B h_C & c_B \\ d_B & d_B b_C + e_B + f_B h_C & f_B \\ g_B & g_B b_C + h_B + \ell_B h_C & \ell_B \end{pmatrix}$$

and

$$CB = \begin{pmatrix} 1 & b_C & 0 \\ 0 & 1 & 0 \\ 0 & h_C & 1 \end{pmatrix} \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix} = \begin{pmatrix} a_B + d_B b_C & b_B + e_B b_C & c_B + f_B b_C \\ d_B & e_B & f_B \\ d_B h_C + g_B & e_B h_C + h_B & f_B h_C + \ell_B \end{pmatrix}.$$

Since $BC = CB$, we have $d_B b_C = 0$, $d_B h_C = 0$, $f_B b_C = 0$, and $f_B h_C = 0$. By the supposition $h_C \neq 0$, we further deduce that $d_B = f_B = 0$. Then, B should be

$$B = \begin{pmatrix} a_B & b_B & c_B \\ 0 & e_B & 0 \\ g_B & h_B & \ell_B \end{pmatrix}.$$

Note that we also have

$$a_B b_C + c_B h_C = e_B b_C \text{ and } g_B b_C + \ell_B h_C = e_B h_C \tag{8}$$

by the equality $BC = CB$.

The characteristic polynomial of B is $P(x) = -x^3 + \mathrm{tr}(B)x^2 - (a_B e_B + a_B \ell_B + e_B \ell_B - c_B g_B)x + \det(B)$ which has roots $\lambda = e_B$ and $\lambda = \frac{1}{2}(a_B + \ell_B \pm \sqrt{(a_B - \ell_B)^2 + 4 c_B g_B})$. We know that the only eigenvalue of B is 1 by Lemma 3 and therefore, we have $a_B = e_B = \ell_B = 1$ and $c_B g_B = 0$.

Moreover, it follows from Equation (8) that $c_B = 0$ and $g_B b_C = 0$. Note that $g_B \neq 0$ because otherwise the matrix B commutes with A. Finally, we consider

$$\begin{aligned} BD &= \begin{pmatrix} 1 & b_B & 0 \\ 0 & 1 & 0 \\ g_B & h_B & 1 \end{pmatrix} \begin{pmatrix} 1 & b_D & c_D \\ 0 & 1 & 0 \\ 0 & h_D & 1 \end{pmatrix} = \begin{pmatrix} 1 & b_B + b_D & c_D \\ 0 & 1 & 0 \\ g_B & g_B b_D + h_B + h_D & g_B c_D + 1 \end{pmatrix} \\ &= \begin{pmatrix} c_D g_B + 1 & b_B + b_D + c_D h_B & c_D \\ 0 & 1 & 0 \\ g_B & h_B + h_D & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & b_D & c_D \\ 0 & 1 & 0 \\ 0 & h_D & 1 \end{pmatrix} \begin{pmatrix} 1 & b_B & 0 \\ 0 & 1 & 0 \\ g_B & h_B & 1 \end{pmatrix} = DB. \end{aligned}$$

It is easy to see that $b_D = c_D = 0$ and then D commutes with C. Therefore, we have a contradiction.

Case 2 ($c_C \neq 0$ **and** $h_C = 0$)**.** Consider the matrix B which commutes with C as follows:

$$\begin{aligned} BC &= \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix} \begin{pmatrix} 1 & b_C & c_C \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_B & a_B b_C + b_B & a_B c_C + c_B \\ d_B & d_B b_C + e_B & d_B c_C + f_B \\ g_B & g_B b_C + h_B & g_B c_C + \ell_B \end{pmatrix} \\ &= \begin{pmatrix} a_B + d_B b_C + g_B c_C & b_B + e_B b_C + h_B c_C & c_B + f_B b_C + \ell_B c_C \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix} \\ &= \begin{pmatrix} 1 & b_C & c_C \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_B & b_B & c_B \\ d_B & e_B & f_B \\ g_B & h_B & \ell_B \end{pmatrix} = CB. \end{aligned}$$

By the equivalence, we have $d_B b_C = 0$, $g_B b_C = 0$, $g_B c_C = 0$, and $d_B c_C = 0$. By the supposition $c_C \neq 0$, we further deduce that $d_B = g_B = 0$. Then, B should be of the following form:

$$B = \begin{pmatrix} a_B & b_B & c_B \\ 0 & e_B & f_B \\ 0 & h_B & \ell_B \end{pmatrix}.$$

Note that we also have

$$a_B b_C = e_B b_C + h_B c_C \text{ and } a_B c_C = f_B b_C + \ell_B c_C \tag{9}$$

by the equality $BC = CB$.

The characteristic polynomial of B is $P(x) = -x^3 + \mathrm{tr}(B)x^2 - (a_B e_B + a_B \ell_B + e_B \ell_B - f_B h_B)x + \det(B)$ which has roots $\lambda = e_B$ and $\lambda = \frac{1}{2}(e_B + \ell_B \pm \sqrt{(a_B - \ell_B)^2 + 4 f_B h_B})$. We know that the only eigenvalue of B is 1 by Lemma 3 and therefore, we have $a_B = e_B = \ell_B = 1$ and $f_B h_B = 0$.

We can further deduce from Equation (9) that $h_B = 0$ and $f_B b_C = 0$. By a similar argument for the matrices B and D that should commute with each other as in the first case, we have a contradiction.

Case 3 ($c_C \neq 0$ **and** $h_C \neq 0$)**.** It is obvious that c_D and h_D are also non-zero because otherwise C and D should commute. Now consider the matrix B which is commuting with C and D. We can deduce from the relation $BC = CB$ that $d_B = g_B = f_B = 0$ and $a_B = e_B = \ell_B = 1$ since they are eigenvalues of B. Hence,

$$B = \begin{pmatrix} 1 & b_B & c_B \\ 0 & 1 & 0 \\ 0 & h_B & 1 \end{pmatrix}.$$

Now we have $c_C h_B = c_B h_C$ since B and C commute with each other. Note that h_B and c_B are both non-zero since A and B commute if $h_B = c_B = 0$. Let us $\frac{c_C}{h_C} = \frac{c_B}{h_B} = x$. We also have $c_D h_B = c_B h_D$ from the relation $BD = DB$ and have $\frac{c_D}{h_D} = \frac{c_B}{h_B} = x$. From $x = \frac{c_C}{h_C} = \frac{c_D}{h_D}$, we have $c_C h_D = c_D h_C$ which results in the relation $CD = DC$. Therefore, we also have a contradiction.

Since we have examined all possible cases and found contradictions for every case, we can conclude that there is no injective morphism from $\Sigma^* \times \Sigma^*$ into the special linear group $\mathrm{SL}(3, \mathbb{Z})$. ◀

▶ **Theorem 6.** *There is no injective morphism $\varphi : \Sigma^* \times \Sigma^* \to \mathrm{H}(3,\mathbb{C})$ for any alphabet Σ with at least two elements.*

Proof. Assume to the contrary that there is an injective morphism φ from $\Sigma^* \times \Sigma^*$ into $\mathrm{H}(3,\mathbb{C})$. Using the notations and relations of (2), we set $\varphi(a) = A$, $\varphi(b) = B$, $\varphi(c) = C$, $\varphi(d) = D$, $\varphi(e) = E$ for some matrices $A, B, C, D, E \in \mathrm{H}(3,\mathbb{C})$. It is easy to see that two matrices $M, N \in \mathrm{H}(3,\mathbb{C})$ commute if and only if $\vec{v}(M) \times \vec{v}(N) = 0$. Denote $\vec{v}(A) = (a_1, a_2)$ and $\vec{v}(B), \vec{v}(C), \vec{v}(D), \vec{v}(E)$ are denoted analogously. From the relations (2), it follows that

$$a_1c_2 = c_1a_2, \quad a_1d_2 = d_1a_2 \quad b_1c_2 = c_1b_2,, \quad b_1d_2 = d_1b_2, \quad a_1b_2 \neq b_1a_2, \quad c_1d_2 \neq d_1c_2.$$

Observe first, that $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2 \neq 0$. If, say, $a_1 = 0$, then from the first two equations it follows that either $a_2 = 0$ or $c_1 = d_1 = 0$. If $a_2 = 0$, then the first inequality does not hold, since $a_1b_2 = 0 = b_1a_2$, and if $c_1 = d_1 = 0$, then the second inequality does not hold, since $c_1d_2 = 0 = d_1c_2$.

Now, we can solve a_1 from the first two equalities, $\frac{a_2c_1}{c_2} = a_1 = \frac{a_2d_1}{d_2}$. That is, $c_1d_2 = d_1c_2$, which contradicts the last relation and proves our claim. ◀

▶ **Lemma 8.** *Let $S = \langle M_1, M_2, \ldots, M_n\rangle \subseteq \mathrm{H}(3,\mathbb{Q})$ be a finitely generated matrix semigroup. Then, the identity matrix exists in S if there exists a sequence of matrices $M_{i_1}M_{i_2}\cdots M_{i_k}$, where $i_j \in [1,n]$ for all $1 \leq j \leq k$, satisfying the following properties:*

1. *$\psi(M_{i_1}M_{i_2}\cdots M_{i_k}) = (0,0,c)$ for some $c \in \mathbb{Q}$, and*
2. *$\vec{v}(M_{i_{j_1}})$ and $\vec{v}(M_{i_{j_2}})$ are not parallel for some $j_1, j_2 \in [1,k]$.*

Proof. Let $M = M_{i_1}M_{i_2}\cdots M_{i_k}$ and $\psi(M) = (0,0,c)$ for some $c \in \mathbb{Q}$. It is obvious that the identity matrix is in S if $c = 0$. Hence we assume that $c > 0$ as the case of $c < 0$ is symmetric.

Given that M_i is the ith generator and $\psi(M_i) = (a_i, b_i, c_i)$, we have $\sum_{j=1}^{k} a_{i_j} = 0$ and $\sum_{j=1}^{k} b_{i_j} = 0$. Since $c > 0$, the following also holds:

$$c = \sum_{i=1}^{k-1}\sum_{j=1}^{i} a_jb_{i+1} + \sum_{j=1}^{k} c_{i_j} > 0. \tag{3}$$

If the matrix semigroup S in $\mathrm{H}(3,\mathbb{Q})$ has two different matrices M_1 and M_2 such that $\psi(M_1) = (0,0,c_1)$ and $\psi(M_2) = (0,0,c_2)$ and $c_1c_2 < 0$, then the identity matrix should exist in S. Let $\psi(M_1) = (0,0,\frac{p_1}{q_1})$ and $\psi(M_2) = (0,0,\frac{p_2}{q_2})$ where $p_1, q_1, q_2 \in \mathbb{Z}$ are positive and $p_2 \in \mathbb{Z}$ is negative. Then, it is easy to see that the matrix $M_1^{-q_1p_2}M_2^{q_2p_1}$ exists in S such that $\psi(M_1^{-q_1p_2}M_2^{q_2p_1}) = (0,0,0)$.

Now we will prove that if S contains a matrix M such that $\psi(M) = (0,0,c)$ where $c > 0$, then there also exists a matrix M' such that $\psi(M') = (0,0,c')$ where $c' < 0$.

First, we classify the matrices into four types as follows. A matrix with a superdiagonal vector (a,b) is classified as

1. the $(+,+)$-type if $a, b > 0$,
2. the $(+,-)$-type if $a \geq 0$ and $b \leq 0$,
3. the $(-,-)$-type if $a, b \leq 0$, and
4. the $(-,+)$-type if $a < 0$ and $b > 0$.

Let $G = \{M_1, M_2, \ldots, M_n\}$ be the generating set of the matrix semigroup S. Then, $G = G_{(+,+)} \sqcup G_{(+,-)} \sqcup G_{(-,-)} \sqcup G_{(-,+)}$ such that $G_{(\xi_1,\xi_2)}$ is the set of matrices of the (ξ_1,ξ_2)-type where $\xi_1, \xi_2 \in \{+,-\}$.

The main idea of the proof is to generate a matrix M' such that $\psi(M') = (0, 0, c')$ for some $c' < 0$ by using the sequence $M = M_{i_1} M_{i_2} \cdots M_{i_k}$ multiple times. Note that any permutation of the sequence generating the matrix M such that $\psi(M) = (0, 0, c)$ still generates matrices M' such that $\psi(M') = (0, 0, c')$ since the multiplication of matrices changes the front two coordinates in a commutative way. Moreover, we can still obtain matrices M'' such that $\psi(M'') = (0, 0, c'')$ for some $c'' \in \mathbb{Q}$ if we shuffle two different permutations of the sequence by the same reason.

We illustrate the idea with the following example. Let $\{M_i \mid 1 \leq i \leq 4\} \subseteq G_{(+,+)}$, $\{M_i \mid 5 \leq i \leq 7\} \subseteq G_{(+,-)}$, $\{M_i \mid 8 \leq i \leq 9\} \subseteq G_{(-,-)}$, and $\{M_i \mid 10 \leq i \leq 13\} \subseteq G_{(-,+)}$. Then, assume that $M_1 M_2 \cdots M_{13} = \left(\begin{smallmatrix} 1 & 0 & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, where x is computed by (3). As we mentioned above, x changes if we change the order of multiplications. See Fig. 2 for example. In this example, we first multiply $(+,+)$-type matrices and accumulate the values in the superdiagonal entries since these matrices have positive values in the entries. Indeed, the blue dotted area implies the value we add to the upper-right corner by multiplying such matrices. Then, we multiply $(+,-)$-type matrices and still increase the 'a'-value. The 'b'-values in $(+,-)$-type matrices are negative thus, the red lined area is subtracted from the upper-right corner. We still subtract by multiplying $(-,-)$-type matrices since the accumulated 'a'-value is still positive and 'b'-values are negative. Then, we finish the multiplication by adding exactly the last blue dotted area to the upper-right corner. It is easy to see that the total subtracted value is larger than the total added value.

However, we cannot guarantee that x is negative since $\sum_{i=1}^{13} c_i$ could be larger than the contribution from the superdiagonal entries. This is why we need to copy the sequence of matrices generating the matrix corresponding to the triple $(0, 0, c)$ for some $c \in \mathbb{Z}$. In Fig. 3, we describe an example where we duplicate the sequence eight times and shuffle and permute them in order to minimize the value in the upper-right corner. Now the lengths of both axes are m ($m = 8$ in Fig. 3) times larger than before and it follows that the area also grows quadratically in m. Since the summation $m \cdot \sum_{i=1}^{13} c_i$ grows linearly in m, we have $x < 0$ when m is large enough.

From the sequence $M_{i_1} M_{i_2} \cdots M_{i_k}$, we obtain four multisets $S_{(\xi_1,\xi_2)}$, where $\xi_1, \xi_2 \in \{+,-\}$, such that each multiset $S_{(\xi_1,\xi_2)}$ contains the matrices that appear in the sequence and belong to the set $G_{(\xi_1,\xi_2)}$. For instance, the multiset $S_{(+,+)}$ has two identical matrices corresponding to $(5, 4, 2)$ which obviously belongs to $G_{(+,+)}$ if the matrix appears twice in the sequence $M_{i_1} M_{i_2} \cdots M_{i_k}$ since we allow any multiset to have multiple identical elements.

For each $\xi_1, \xi_2 \in \{+,-\}$, let us define $a_{(\xi_1,\xi_2)}, b_{(\xi_1,\xi_2)}, c_{(\xi_1,\xi_2)}$ such that

$$(a_{(\xi_1,\xi_2)}, b_{(\xi_1,\xi_2)}, c_{(\xi_1,\xi_2)}) = \sum_{M \in S_{(\xi_1,\xi_2)}} \psi(M).$$

In other words, $a_{(\xi_1,\xi_2)}$ ($b_{(\xi_1,\xi_2)}$ and $c_{(\xi_1,\xi_2)}$, respectively) is the sum of the values in the 'a' ('b' and 'c', respectively) coordinate from the matrices in the multiset $S_{(\xi_1,\xi_2)}$.

Now consider a permutation of the sequence $M_{i_1} M_{i_2} \cdots M_{i_k}$, where the first part of the sequence only consists of the $(+,+)$-type matrices, the second part only consists of the $(+,-)$-type matrices, the third part only consists of the $(-,-)$-type matrices, and finally the last part only consists of the $(-,+)$-type matrices.

Let us denote by $M_{(+,+)}$ the matrix which results from the multiplication of the first part. Then, the matrix $M_{(+,+)}$ will be bounded from above by the following matrix on the

right-hand side:

$$\prod_{M \in S_{(+,+)}} M = \begin{pmatrix} 1 & a_{(+,+)} & x \\ 0 & 1 & b_{(+,+)} \\ 0 & 0 & 1 \end{pmatrix} < \begin{pmatrix} 1 & a_{(+,+)} & c_{(+,+)} + a_{(+,+)}b_{(+,+)} \\ 0 & 1 & b_{(+,+)} \\ 0 & 0 & 1 \end{pmatrix}, \tag{4}$$

where x is computed by (3). Let us define $M_{(+,-)}$, $M_{(-,-)}$ and $M_{(-,+)}$ analogously.

Now we claim that there exists an integer $m > 0$ such that $M^m_{(+,+)}M^m_{(+,-)}M^m_{(-,-)}M^m_{(-,+)}$ corresponds to the triple $(0, 0, c')$ for some $c' < 0$.

Let us first consider the first part $M^m_{(+,+)}$. It follows from (4) that $\psi(M^m_{(+,+)}) = (ma_{(+,+)}, mb_{(+,+)}, mc_{(+,+)} + z_1)$, where z_1 is bounded from above by the following value:

$$z_1 < \sum_{i=1}^{m} i|a_{(+,+)}||b_{(+,+)}| = \frac{m(m+1)}{2} \cdot |a_{(+,+)}||b_{(+,+)}| = z'_1.$$

Now we multiply $M^m_{(+,+)}$ by the second part $M^m_{(+,-)}$. Then, the resulting matrix corresponds to $(m(a_{(+,+)} + a_{(+,-)}), m(b_{(+,+)} + b_{(+,-)}), m(c_{(+,+)} + c_{(+,-)}) + z_1 - z_2)$, where z_2 is bounded from below by the following value:

$$z_2 > m^2|a_{(+,+)}||b_{(+,-)}| + \sum_{i=1}^{m-1} i|a_{(+,-)}||b_{(+,-)}| = z'_2.$$

Similarly, we compute bounds of z_3 and z_4 that are added to the third component of the triple as a result of multiplying $M^m_{(-,-)}$ and $M^m_{(-,+)}$ as follows:

$$z_3 > m^2|a_{(-,+)}||b_{(-,-)}| + \sum_{i=1}^{m-1} i|a_{(-,-)}||b_{(-,-)}| = z'_3$$

and

$$z_4 < \sum_{i=1}^{m} i|a_{(-,+)}||b_{(-,+)}| = \frac{m(m+1)}{2} \cdot |a_{(-,+)}||b_{(-,+)}| = z'_4.$$

It is easy to see that

$$\psi(M^m_{(+,+)}M^m_{(+,-)}M^m_{(-,-)}M^m_{(-,+)}) = (0, 0, mc + z_1 - z_2 - z_3 + z_4).$$

From the above inequalities we obtain the following inequality:

$$z_1 - z_2 - z_3 + z_4 < z'_1 - z'_2 - z'_3 + z'_4 = z.$$

Now we claim that z can be represented as a quadratic equation of m using the above results and then the coefficient of m^2 is always negative if $S_{(\xi_1,\xi_2)} \neq \emptyset$ for all $\xi_1, \xi_2 \in \{+, -\}$.

The coefficient of m^2 in $z'_1 + z'_4$ is

$$\frac{|a_{(+,+)}||b_{(+,+)}| + |a_{(-,+)}||b_{(-,+)}|}{2}$$

and in $z'_2 + z'_3$ is

$$\frac{|a_{(+,-)}||b_{(+,-)}| + |a_{(-,-)}||b_{(-,-)}|}{2} + |a_{(+,+)}||b_{(+,-)}| + |a_{(-,+)}||b_{(-,-)}|.$$

Let $|a_{(+,+)}|+|a_{(+,-)}| = |a_{(-,+)}|+|a_{(-,-)}| = a'$ and $|b_{(+,+)}|+|b_{(-,+)}| = |b_{(+,-)}|+|b_{(-,-)}| = b'$. Then, the coefficient of m^2 in z can be written as

$$-a'b' + \frac{|a_{(+,+)}||b_{(+,+)}| + |a_{(-,+)}||b_{(-,+)}| + |a_{(-,-)}||b_{(-,-)}| + |a_{(+,-)}||b_{(+,-)}|}{2}. \tag{5}$$

Now we will prove that the coefficient of m^2 in (5) is always negative. First, it is easy to see that the following inequalities always hold:

$$|a_{(+,+)}||b_{(+,+)}| + |a_{(-,+)}||b_{(-,+)}| \leq \max\{|a_{(+,+)}|, |a_{(-,+)}|\}b' \leq a'b'$$

and

$$|a_{(-,-)}||b_{(-,-)}| + |a_{(+,-)}||b_{(+,-)}| \leq \max\{|a_{(-,-)}|, |a_{(+,-)}|\}b' \leq a'b'.$$

Moreover, $|a_{(+,+)}||b_{(+,+)}| + |a_{(-,+)}||b_{(-,+)}|$ can be maximized to $a'b'$ only when $|a_{(+,+)}| = |a_{(-,+)}| = a'$. However, as soon as we have $|a_{(+,+)}| = |a_{(-,+)}| = a'$, $|a_{(-,-)}||b_{(-,-)}| + |a_{(+,-)}||b_{(+,-)}|$ becomes 0 since we are assuming $a_{(+,+)}+a_{(+,-)} = a'$ and $b_{(+,+)}+b_{(-,+)} = b'$. We can also see that the case of maximizing $|a_{(-,-)}||b_{(-,-)}| + |a_{(+,-)}||b_{(+,-)}|$ is symmetric. Therefore, we can prove that the coefficient of the highest power in z is always negative if none of the subsets from $S_{(+,+)}$. $S_{(+,-)}$, $S_{(-,-)}$, and $S_{(-,+)}$ is empty.

There are some subcases where some of subsets from $S_{(+,+)}$, $S_{(-,+)}$, $S_{(+,-)}$, and $S_{(-,-)}$ is empty. Fig. 4 shows the cases when one of subsets from $S_{(+,+)}$, $S_{(-,+)}$, $S_{(+,-)}$, and $S_{(-,-)}$ is empty. It is easily seen that we have larger area for the negative contribution in any subcase. We also formally prove that the coefficient of m^2 in z should be negative when only one of the subsets from $S_{(+,+)}$. $S_{(+,-)}$, $S_{(-,-)}$, and $S_{(-,+)}$ is empty as follows:

1. The case of $S_{(+,+)} = \emptyset$: Note that $|a_{(+,-)}| = a'$ and $|b_{(-,+)}| = b'$ since $|a_{(+,+)}| = |b_{(+,+)}| = 0$ by $S_{(+,+)} = \emptyset$ being empty. Then, the coefficient of m^2 becomes $-a'b' + \frac{|a_{(-,+)}|b'+|a_{(-,-)}||b_{(-,-)}|+a'|b_{(+,-)}|}{2}$. We can see that the coefficient can be at most 0 since $|a_{(-,+)}|b'$ and $|a_{(-,-)}||b_{(-,-)}| + a'|b_{(+,-)}|$ can be maximized to $a'b'$. If we maximize $|a_{(-,+)}|b'$ by setting $|a_{(-,+)}| = a'$, then $|a_{(-,-)}|$ should be 0 since $|a_{(+,+)}| + |a_{(-,+)}| = a'$. Then, $|a_{(-,-)}||b_{(-,-)}| + a'|b_{(+,-)}|$ can be $a'b'$ only when $|b_{(+,-)}| = b'$. This leads to the set $S_{(-,-)}$ being empty since we have $|a_{(-,-)}| = 0$ and $|b_{(-,-)}| = 0$ and therefore, we have a contradiction.
2. The case of $S_{(+,-)} = \emptyset$, $S_{(-,-)} = \emptyset$, or $S_{(-,+)} = \emptyset$: We can prove the remaining cases by the similar argument as above.

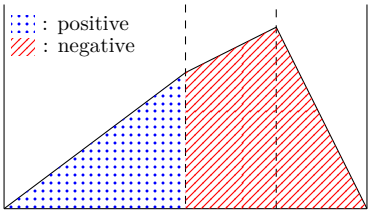
Lastly, it remains to consider the cases where two of the subsets are empty. Note that we do not consider the cases where three of the subsets are empty because the sum of a's and b's cannot be both zero in such cases. Here we assume one of $S_{(+,+)}$ and $S_{(-,-)}$ should contain two matrices whose superdiagonal vectors are not parallel by the statement of this lemma. Then, we can always make the negative contribution larger by using matrices with different superdiagonal vectors. See Fig. 5 for example. More formally, we consider the two cases as follows:

1. The case of $S_{(+,+)} = \emptyset$ and $S_{(-,-)} = \emptyset$: Without loss of generality, assume that $S_{(-,+)}$ contains two matrices M_1 and M_2 with non-parallel superdiagonal vectors. Let $\vec{v}(M_1) = (a_1, b_1)$ and $\vec{v}(M_2) = (a_2, b_2)$ be superdiagonal vectors for M_1 and M_2, respectively, such that $|\frac{a_1}{b_1}| > |\frac{a_2}{b_2}|$. To simplify the proof, we assume the set $S_{(+,-)}$ only uses one matrix M_3, where $\vec{v}(M_3) = (a_3, b_3)$, to generate a matrix with a zero superdiagonal

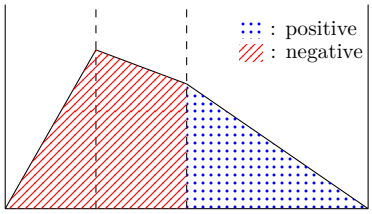
vector. This implies that $a_1x + a_2y + a_3 = 0$ and $b_1x + b_2y + b_3 = 0$ for some $x, y \in \mathbb{Q}$. Here the idea is that we first multiply the matrix M_1 and then multiply M_2 later. For instance, we first multiply M_1^m and then M_2^m. Then, the coefficient of the highest power in z becomes $\frac{-a'b'+2|a_2||b_1|+|a_1||b_1|+|a_2||b_2|}{2}$. Since $a' = |a_1| + |a_2|$ and $b' = |b_1| + |b_2|$, the coefficient of m^2 is now $\frac{|a_2||b_1|-|a_1||b_2|}{2}$. By the supposition $|\frac{a_1}{b_1}| > |\frac{a_2}{b_2}|$, we prove that the coefficient of the highest power in z is always negative.

2. The case of $S_{(+,-)} = \emptyset$ and $S_{(-,+)} = \emptyset$: We can prove this case by the similar argument as above.

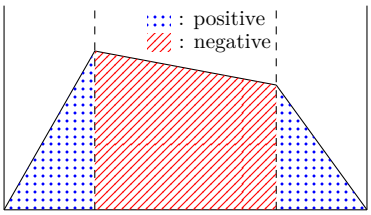
(a) The case of $S_{(-,+)}$ being empty



(b) The case of $S_{(+,+)}$ being empty



(c) The case of $S_{(+,-)}$ being empty



(d) The case of $S_{(-,-)}$ being empty

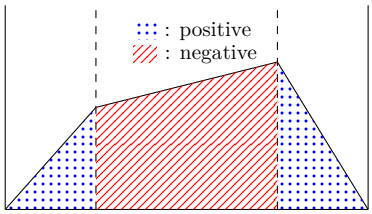
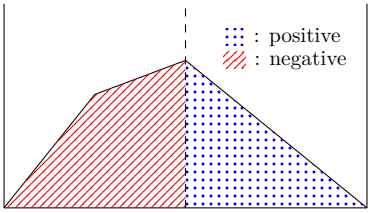


Figure 4 Subcases where one of the subsets from $S_{(+,+)}$, $S_{(-,+)}$, $S_{(+,-)}$, and $S_{(-,-)}$ is empty

(a) When $S_{(+,+)}$ and $S_{(-,-)}$ are empty



(b) When $S_{(-,+)}$ and $S_{(+,-)}$ are empty

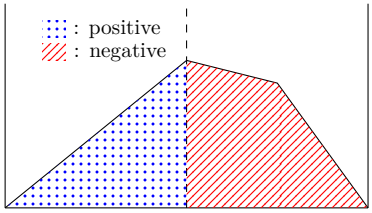


Figure 5 Subcases where two of the subsets from $S_{(+,+)}$, $S_{(-,+)}$, $S_{(+,-)}$, and $S_{(-,-)}$ are empty

Since we have proven that it is always possible to construct a matrix M' such that $\psi(M') = (0, 0, c')$ for some $c' < 0$ in any case, we complete the proof. ◀

▶ **Example 9.** We illustrate Lemma 8. Consider a semigroup S generated by matrices

$$\begin{pmatrix} 1 & -4 & 20 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 20 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 20 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 20 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix}.$$

A simple calculation shows that a product of the four matrices (in any order) is a matrix M such that $\psi(M) = (0, 0, 80 + x)$ for some $x \in \mathbb{Z}$. Our goal, is to minimize x by multiplying

the matrices in different order. Denote the given matrices by $M_{(+,+)} = \left(\begin{smallmatrix} 1 & 2 & 20 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, $M_{(+,-)} = \left(\begin{smallmatrix} 1 & 3 & 20 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, $M_{(-,-)} = \left(\begin{smallmatrix} 1 & -4 & 20 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ and $M_{(-,+)} = \left(\begin{smallmatrix} 1 & -1 & 20 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ and

$$N_1 = M_{(+,+)}M_{(+,-)}M_{(-,-)}M_{(-,+)} = \left(\begin{smallmatrix} 1 & 0 & 47 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right).$$

That is, $x = -33$. By considering several copies of the product, we can have a negative value in the top right corner. Indeed, consider the product of 16 matrices

$$N_2 = M^4_{(+,+)}M^4_{(+,-)}M^4_{(-,-)}M^4_{(-,+)} = \left(\begin{smallmatrix} 1 & 0 & -22 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right).$$

Since, we have a matrix with negative value in the top corner, the identity matrix can be generated for example by the product $N_1^{22}N_2^{47}$.